

SEQUENCE LISTING



(1) GENERAL INFORMATION:

- (i) APPLICANT: Ingham, Phillip W.
McMahon, Andrew P.
Tabin, Clifford J.
- (ii) TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
Proteins and Uses Related Thereto
- (iii) NUMBER OF SEQUENCES: 48
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: FOLEY, HOAG & ELIOT LLP
(B) STREET: One Post Office Square
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02109-2170
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/954,771
(B) FILING DATE: 20-OCT-1997
- E¹
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/462,386
(B) FILING DATE: 05-JUN-1995
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/435,093
(B) FILING DATE: 04-MAY-1995
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/356,060
(B) FILING DATE: 14-DEC-1994
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/176,427
(B) FILING DATE: 30-DEC-1993
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Vincent, Matthew P.
(B) REGISTRATION NUMBER: 36,709
(C) REFERENCE/DOCKET NUMBER: HMV-006.11
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617-832-1000
(B) TELEFAX: 617-832-7000

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1277 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1275

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GTC GAA ATG CTG CTG TTG ACA AGA ATT CTC TTG GTG GGC TTC ATC	48
Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile	
1 5 10 15	
TGC GCT CTT TTA GTC TCC TCT GGG CTG ACT TGT GGA CCA GGC AGG GGC	96
Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly	
20 25 30	
ATT GGA AAA AGG AGG CAC CCC AAA AAG CTG ACC CCG TTA GCC TAT AAG	144
Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys	
35 40 45	
CAG TTT ATT CCC AAT GTG GCA GAG AAG ACC CTA GGG GCC AGT GGA AGA	192
Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg	
50 55 60	
TAT GAA GGG AAG ATC ACA AGA AAC TCC GAG AGA TTT AAA GAA CTA ACC	240
Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr	
65 70 75 80	
CCA AAT TAC AAC CCT GAC ATT ATT TTT AAG GAT GAA GAG AAC ACG GGA	288
Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly	
85 90 95	
GCT GAC AGA CTG ATG ACT CAG CGC TGC AAG GAC AAG CTG AAT GCC CTG	336
Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu	
100 105 110	
GCG ATC TCG GTG ATG AAC CAG TGG CCC GGG GTG AAG CTG CGG GTG ACC	384
Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr	
115 120 125	
GAG GGC TGG GAC GAG GAT GGC CAT CAC TCC GAG GAA TCG CTG CAC TAC	432
Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr	
130 135 140	
GAG GGT CGC GCC GTG GAC ATC ACC ACG TCG GAT CGG GAC CGC AGC AAG	480
Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys	
145 150 155 160	

E1
Cont

TAC GGA ATG CTG GCC CGC CTC GCC GTC GAG GCC GGC TTC GAC TGG GTC	528
Tyr Gly Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val	
165 170 175	
TAC TAC GAG TCC AAG GCG CAC ATC CAC TGC TCC GTC AAA GCA GAA AAC	576
Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn	
180 185 190	
TCA GTG GCA GCG AAA TCA GGA GGC TGC TTC CCT GGC TCA GCC ACA GTG	624
Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val	
195 200 205	
CAC CTG GAG CAT GGA GGC ACC AAG CTG GTG AAG GAC CTG AGC CCT GGG	672
His Leu Glu His Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly	
210 215 220	
GAC CGC GTG CTG GCT GCT GAC GCG GAC GGC CGG CTG CTC TAC AGT GAC	720
Asp Arg Val Leu Ala Ala Asp Ala Asp Gly Arg Leu Leu Tyr Ser Asp	
225 230 235 240	
TTC CTC ACC TTC CTC GAC CGG ATG GAC AGC TCC CGA AAG CTC TTC TAC	768
Phe Leu Thr Phe Leu Asp Arg Met Asp Ser Ser Arg Lys Leu Phe Tyr	
245 250 255	
GTC ATC GAG ACG CGG CAG CCC CGG GCC CGG CTG CTA CTG ACG GCG GCC	816
Val Ile Glu Thr Arg Gln Pro Arg Ala Arg Leu Leu Leu Thr Ala Ala	
260 265 270	
CAC CTG CTC TTT GTG GCC CCC CAG CAC AAC CAG TCG GAG GCC ACA GGG	864
His Leu Leu Phe Val Ala Pro Gln His Asn Gln Ser Glu Ala Thr Gly	
275 280 285	
TCC ACC AGT GGC CAG GCG CTC TTC GCC AGC AAC GTG AAG CCT GGC CAA	912
Ser Thr Ser Gly Gln Ala Leu Phe Ala Ser Asn Val Lys Pro Gly Gln	
290 295 300	
CGT GTC TAT GTG CTG GGC GAG GGC GGG CAG CAG CTG CTG CCG GCG TCT	960
Arg Val Tyr Val Leu Glu Gly Gly Gln Gln Leu Leu Pro Ala Ser	
305 310 315 320	
GTC CAC AGC GTC TCA TTG CGG GAG GAG GCG TCC GGA GCC TAC GCC CCA	1008
Val His Ser Val Ser Leu Arg Glu Glu Ala Ser Gly Ala Tyr Ala Pro	
325 330 335	
CTC ACC GCC CAG GGC ACC ATC CTC ATC AAC CGG GTG TTG GCC TCC TGC	1056
Leu Thr Ala Gln Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys	
340 345 350	
TAC GCC GTC ATC GAG GAG CAC AGT TGG GCC CAT TGG GCC TTC GCA CCA	1104
Tyr Ala Val Ile Glu Glu His Ser Trp Ala His Trp Ala Phe Ala Pro	
355 360 365	
TTC CGC TTG GCT CAG GGG CTG CTG GCC GCC CTC TGC CCA GAT GGG GCC	1152
Phe Arg Leu Ala Gln Gly Leu Leu Ala Ala Leu Cys Pro Asp Gly Ala	
370 375 380	

E1
cont

ATC CCT ACT GCC GCC ACC ACC ACC ACT GGC ATC CAT TGG TAC TCA CGG	1200
Ile Pro Thr Ala Ala Thr Thr Thr Thr Gly Ile His Trp Tyr Ser Arg	
385 390 395 400	
CTC CTC TAC CGC ATC GGC AGC TGG GTG CTG GAT GGT GAC GCG CTG CAT	1248
Leu Leu Tyr Arg Ile Gly Ser Trp Val Leu Asp Gly Asp Ala Leu His	
405 410 415	
CCG CTG GGC ATG GTG GCA CCG GCC AGC TG	1277
Pro Leu Gly Met Val Ala Pro Ala Ser	
420 425	

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1190 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1188

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATG GCT CTG CCG GCC AGT CTG TTG CCC CTG TGC TGC TTG GCA CTC TTG	48
Met Ala Leu Pro Ala Ser Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu	
1 5 10 15	
GCA CTA TCT GCC CAG AGC TGC GGG CCG GGC CGA GGA CCG GTT GGC CGG	96
Ala Leu Ser Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg	
20 25 30	
CGG CGT TAT GTG CGC AAG CAA CTT GTG CCT CTG CTA TAC AAG CAG TTT	144
Arg Arg Tyr Val Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe	
35 40 45	
GTG CCC AGT ATG CCC GAG CGG ACC CTG GGC GCG AGT GGG CCA GCG GAG	192
Val Pro Ser Met Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu	
50 55 60	
GGG AGG GTA ACA AGG GGG TCG GAG CGC TTC CGG GAC CTC GTA CCC AAC	240
Gly Arg Val Thr Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn	
65 70 75 80	
TAC AAC CCC GAC ATA ATC TTC AAG GAT GAG GAG AAC AGC GGC GCA GAC	288
Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp	
85 90 95	
CGC CTG ATG ACA GAG CGT TGC AAA GAG CGG GTG AAC GCT CTA GCC ATC	336
Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile	
100 105 110	

E1
Cont

GCG	GTG	ATG	AAC	ATG	TGG	CCC	GGA	GTA	CGC	CTA	CGT	GTG	ACT	GAA	GGC	384
Ala	Val	Met	Asn	Met	Trp	Pro	Gly	Val	Arg	Leu	Arg	Val	Thr	Glu	Gly	
		115					120					125				
TGG	GAC	GAG	GAC	GGC	CAC	CAC	GCA	CAG	GAT	TCA	CTC	CAC	TAC	GAA	GGC	432
Trp	Asp	Glu	Asp	Gly	His	His	Ala	Gln	Asp	Ser	Leu	His	Tyr	Glu	Gly	
	130						135					140				
CGT	GCC	TTG	GAC	ATC	ACC	ACG	TCT	GAC	CGT	GAC	CGT	AAT	AAG	TAT	GGT	480
Arg	Ala	Leu	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Arg	Asn	Lys	Tyr	Gly	
	145				150					155					160	
TTG	TTG	GCG	CGC	CTA	GCT	GTG	GAA	GCC	GGA	TTC	GAC	TGG	GTC	TAC	TAC	528
Leu	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	
				165					170					175		
GAG	TCC	CGC	AAC	CAC	ATC	CAC	GTA	TCG	GTC	AAA	GCT	GAT	AAC	TCA	CTG	576
Glu	Ser	Arg	Asn	His	Ile	His	Val	Ser	Val	Lys	Ala	Asp	Asn	Ser	Leu	
			180					185					190			
GCG	GTC	CGA	GCC	GGA	GGC	TGC	TTT	CCG	GGA	AAT	GCC	ACG	GTG	CGC	TTG	624
Ala	Val	Arg	Ala	Gly	Gly	Cys	Phe	Pro	Gly	Asn	Ala	Thr	Val	Arg	Leu	
		195					200					205				
CGG	AGC	GGC	GAA	CGG	AAG	GGG	CTG	AGG	GAA	CTA	CAT	CGT	GGT	GAC	TGG	672
Arg	Ser	Gly	Glu	Arg	Lys	Gly	Leu	Arg	Glu	Leu	His	Arg	Gly	Asp	Trp	
	210					215					220					
GTA	CTG	GCC	GCT	GAT	GCA	GCG	GGC	CGA	GTG	GTA	CCC	ACG	CCA	GTG	CTG	720
Val	Leu	Ala	Ala	Asp	Ala	Ala	Gly	Arg	Val	Val	Pro	Thr	Pro	Val	Leu	
	225				230					235					240	
CTC	TTC	CTG	GAC	CGG	GAT	CTG	CAG	CGC	CGC	GCC	TCG	TTC	GTG	GCT	GTG	768
Leu	Phe	Leu	Asp	Arg	Asp	Leu	Gln	Arg	Arg	Ala	Ser	Phe	Val	Ala	Val	
				245					250					255		
GAG	ACC	GAG	CGG	CCT	CCG	CGC	AAA	CTG	TTG	CTC	ACA	CCC	TGG	CAT	CTG	816
Glu	Thr	Glu	Arg	Pro	Pro	Arg	Lys	Leu	Leu	Leu	Thr	Pro	Trp	His	Leu	
			260					265					270			
GTG	TTC	GCT	GCT	CGC	GGG	CCA	GCG	CCT	GCT	CCA	GGT	GAC	TTT	GCA	CCG	864
Val	Phe	Ala	Ala	Arg	Gly	Pro	Ala	Pro	Ala	Pro	Gly	Asp	Phe	Ala	Pro	
		275					280					285				
GTG	TTC	GCG	CGC	CGC	TTA	CGT	GCT	GGC	GAC	TCG	GTG	CTG	GCT	CCC	GGC	912
Val	Phe	Ala	Arg	Arg	Leu	Arg	Ala	Gly	Asp	Ser	Val	Leu	Ala	Pro	Gly	
		290				295					300					
GGG	GAC	GCG	CTC	CAG	CCG	GCG	CGC	GTA	GCC	CGC	GTG	GCG	CGC	GAG	GAA	960
Gly	Asp	Ala	Leu	Gln	Pro	Ala	Arg	Val	Ala	Arg	Val	Ala	Arg	Glu	Glu	
	305				310					315					320	
GCC	GTG	GGC	GTG	TTC	GCA	CCG	CTC	ACT	GCG	CAC	GGG	ACG	CTG	CTG	GTC	1008
Ala	Val	Gly	Val	Phe	Ala	Pro	Leu	Thr	Ala	His	Gly	Thr	Leu	Leu	Val	
				325					330					335		

E1
Cont

AAC GAC GTC CTC GCC TCC TGC TAC GCG GTT CTA GAG AGT CAC CAG TGG	1056
Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp	
340 345 350	
GCC CAC CGC GCC TTC GCC CCT TTG CGG CTG CTG CAC GCG CTC GGG GCT	1104
Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala	
355 360 365	
CTG CTC CCT GGG GGT GCA GTC CAG CCG ACT GGC ATG CAT TGG TAC TCT	1152
Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser	
370 375 380	
CGC CTC CTT TAC CGC TTG GCC GAG GAG TTA ATG GGC TG	1190
Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Met Gly	
385 390 395	

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1281 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1233

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG TCT CCC GCC TGG CTC CGG CCC CGA CTG CGG TTC TGT CTG TTC CTG	48
Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu	
1 5 10 15	
CTG CTG CTG CTT CTG GTG CCG GCG GCG CGG GGC TGC GGG CCG GGC CGG	96
Leu Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg	
20 25 30	
GTG GTG GGC AGC CGC CGG AGG CCG CCT CGC AAG CTC GTG CCT CTT GCC	144
Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala	
35 40 45	
TAC AAG CAG TTC AGC CCC AAC GTG CCG GAG AAG ACC CTG GGC GCC AGC	192
Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser	
50 55 60	
GGG CGC TAC GAA GGC AAG ATC GCG CGC AGC TCT GAG CGC TTC AAA GAG	240
Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu	
65 70 75 80	
CTC ACC CCC AAC TAC AAT CCC GAC ATC ATC TTC AAG GAC GAG GAG AAC	288
Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn	
85 90 95	

E1
Cont

ACG GGT GCC GAC CGC CTC ATG ACC CAG CGC TGC AAG GAC CGT CTG AAC	336
Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn	
100 105 110	
TCA CTG GCC ATC TCT GTC ATG AAC CAG TGG CCT GGT GTG AAA CTG CGG	384
Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg	
115 120 125	
GTG ACC GAA GGC CGG GAT GAA GAT GGC CAT CAC TCA GAG GAG TCT TTA	432
Val Thr Glu Gly Arg Asp Glu Asp Gly His His Ser Glu Glu Ser Leu	
130 135 140	
CAC TAT GAG GGC CGC GCG GTG GAT ATC ACC ACC TCA GAC CGT GAC CGA	480
His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg	
145 150 155 160	
AAT AAG TAT GGA CTG CTG GCG CGC TTA GCA GTG GAG GCC GGC TTC GAC	528
Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp	
165 170 175	
TGG GTG TAT TAC GAG TCC AAG GCC CAC GTG CAT TGC TCT GTC AAG TCT	576
Trp Val Tyr Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser	
180 185 190	
GAG CAT TCG GCC GCT GCC AAG ACA GGT GGC TGC TTT CCT GCC GGA GCC	624
Glu His Ser Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala	
195 200 205	
CAG GTG CGC CTA GAG AAC GGG GAG CGT GTG GCC CTG TCA GCT GTA AAG	672
Gln Val Arg Leu Glu Asn Gly Glu Arg Val Ala Leu Ser Ala Val Lys	
210 215 220	
CCA GGA GAC CGG GTG CTG GCC ATG GGG GAG GAT GGG ACC CCC ACC TTC	720
Pro Gly Asp Arg Val Leu Ala Met Gly Glu Asp Gly Thr Pro Thr Phe	
225 230 235 240	
AGT GAT GTG CTT ATT TTC CTG GAC CGC GAG CCA AAC CGG CTG AGA GCT	768
Ser Asp Val Leu Ile Phe Leu Asp Arg Glu Pro Asn Arg Leu Arg Ala	
245 250 255	
TTC CAG GTC ATC GAG ACT CAG GAT CCT CCG CGT CGG CTG GCG CTC ACG	816
Phe Gln Val Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr	
260 265 270	
CCT GCC CAC CTG CTC TTC ATT GCG GAC AAT CAT ACA GAA CCA GCA GCC	864
Pro Ala His Leu Leu Phe Ile Ala Asp Asn His Thr Glu Pro Ala Ala	
275 280 285	
CAC TTC CGG GCC ACA TTT GCC AGC CAT GTG CAA CCA GGC CAA TAT GTG	912
His Phe Arg Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val	
290 295 300	
CTG GTA TCA GGG GTA CCA GGC CTC CAG CCT GCT CGG GTG GCA GCT GTC	960
Leu Val Ser Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val	
305 310 315 320	

E1
Cont.

TCC ACC CAC GTG GCC CTT GGG TCC TAT GCT CCT CTC ACA AGG CAT GGG	1008
Ser Thr His Val Ala Leu Gly Ser Tyr Ala Pro Leu Thr Arg His Gly	
325 330 335	
ACA CTT GTG GTG GAG GAT GTG GTG GCC TCC TGC TTT GCA GCT GTG GCT	1056
Thr Leu Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala	
340 345 350	
GAC CAC CAT CTG GCT CAG TTG GCC TTC TGG CCC CTG CGA CTG TTT CCC	1104
Asp His His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe Pro	
355 360 365	
AGT TTG GCA TGG GGC AGC TGG ACC CCA AGT GAG GGT GTT CAC TCC TAC	1152
Ser Leu Ala Trp Gly Ser Trp Thr Pro Ser Glu Gly Val His Ser Tyr	
370 375 380	
CCT CAG ATG CTC TAC CGC CTG GGG CGT CTC TTG CTA GAA GAG AGC ACC	1200
Pro Gln Met Leu Tyr Arg Leu Gly Arg Leu Leu Glu Glu Ser Thr	
385 390 395 400	
TTC CAT CCA CTG GGC ATG TCT GGG GCA GGA AGC TGAAGGGACT CTAACCACTG	1253
Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser	
405 410	
CCCTCCTGGA ACTGCTGTGC GTGGATCC	1281

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1313 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1311

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATG CTG CTG CTG CTG GCC AGA TGT TTT CTG GTG ATC CTT GCT TCC TCG	48
Met Leu Leu Leu Leu Ala Arg Cys Phe Leu Val Ile Leu Ala Ser Ser	
1 5 10 15	
CTG CTG GTG TGC CCC GGG CTG GCC TGT GGG CCC GGC AGG GGG TTT GGA	96
Leu Leu Val Cys Pro Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly	
20 25 30	
AAG AGG CGG CAC CCC AAA AAG CTG ACC CCT TTA GCC TAC AAG CAG TTT	144
Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe	
35 40 45	

E1
Cont.

ATT	CCC	AAC	GTA	GCC	GAG	AAG	ACC	CTA	GGG	GCC	AGC	GGC	AGA	TAT	GAA	192
Ile	Pro	Asn	Val	Ala	Glu	Lys	Thr	Leu	Gly	Ala	Ser	Gly	Arg	Tyr	Glu	
	50					55					60					
GGG	AAG	ATC	ACA	AGA	AAC	TCC	GAA	CGA	TTT	AAG	GAA	CTC	ACC	CCC	AAT	240
Gly	Lys	Ile	Thr	Arg	Asn	Ser	Glu	Arg	Phe	Lys	Glu	Leu	Thr	Pro	Asn	
65					70				75						80	
TAC	AAC	CCC	GAC	ATC	ATA	TTT	AAG	GAT	GAG	GAA	AAC	ACG	GGA	GCA	GAC	288
Tyr	Asn	Pro	Asp	Ile	Ile	Phe	Lys	Asp	Glu	Glu	Asn	Thr	Gly	Ala	Asp	
				85					90					95		
CGG	CTG	ATG	ACT	CAG	AGG	TGC	AAA	GAC	AAG	TTA	AAT	GCC	TTG	GCC	ATC	336
Arg	Leu	Met	Thr	Gln	Arg	Cys	Lys	Asp	Lys	Leu	Asn	Ala	Leu	Ala	Ile	
			100					105					110			
TCT	GTG	ATG	AAC	CAG	TGG	CCT	GGA	GTG	AGG	CTG	CGA	GTG	ACC	GAG	GGC	384
Ser	Val	Met	Asn	Gln	Trp	Pro	Gly	Val	Arg	Leu	Arg	Val	Thr	Glu	Gly	
	115						120					125				
TGG	GAT	GAG	GAC	GGC	CAT	CAT	TCA	GAG	GAG	TCT	CTA	CAC	TAT	GAG	GGT	432
Trp	Asp	Glu	Asp	Gly	His	His	Ser	Glu	Glu	Ser	Leu	His	Tyr	Glu	Gly	
130						135					140					
CGA	GCA	GTG	GAC	ATC	ACC	ACG	TCC	GAC	CGG	GAC	CGC	AGC	AAG	TAC	GGC	480
Arg	Ala	Val	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Arg	Ser	Lys	Tyr	Gly	
145					150					155					160	
ATG	CTG	GCT	CGC	CTG	GCT	GTG	GAA	GCA	GGT	TTC	GAC	TGG	GTC	TAC	TAT	528
Met	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	
				165					170					175		
GAA	TCC	AAA	GCT	CAC	ATC	CAC	TGT	TCT	GTG	AAA	GCA	GAG	AAC	TCC	GTG	576
Glu	Ser	Lys	Ala	His	Ile	His	Cys	Ser	Val	Lys	Ala	Glu	Asn	Ser	Val	
		180						185					190			
GCG	GCC	AAA	TCC	GGC	GGC	TGT	TTC	CCG	GGA	TCC	GCC	ACC	GTG	CAC	CTG	624
Ala	Ala	Lys	Ser	Gly	Gly	Cys	Phe	Pro	Gly	Ser	Ala	Thr	Val	His	Leu	
		195					200					205				
GAG	CAG	GGC	GGC	ACC	AAG	CTG	GTG	AAG	GAC	TTA	CGT	CCC	GGA	GAC	CGC	672
Glu	Gln	Gly	Gly	Thr	Lys	Leu	Val	Lys	Asp	Leu	Arg	Pro	Gly	Asp	Arg	
210						215					220					
GTG	CTG	GCG	GCT	GAC	GAC	CAG	GGC	CGG	CTG	CTG	TAC	AGC	GAC	TTC	CTC	720
Val	Leu	Ala	Ala	Asp	Asp	Gln	Gly	Arg	Leu	Leu	Tyr	Ser	Asp	Phe	Leu	
225					230					235					240	
ACC	TTC	CTG	GAC	CGC	GAC	GAA	GGC	GCC	AAG	AAG	GTC	TTC	TAC	GTG	ATC	768
Thr	Phe	Leu	Asp	Arg	Asp	Glu	Gly	Ala	Lys	Lys	Val	Phe	Tyr	Val	Ile	
				245					250					255		
GAG	ACG	CTG	GAG	CCG	CGC	GAG	CGC	CTG	CTG	CTC	ACC	GCC	GCG	CAC	CTG	816
Glu	Thr	Leu	Glu	Pro	Arg	Glu	Arg	Leu	Leu	Leu	Thr	Ala	Ala	His	Leu	
			260					265					270			

E1
Cont.

CTC TTC GTG GCG CCG CAC AAC GAC TCG GGG CCC ACG CCC GGG CCA AGC	864
Leu Phe Val Ala Pro His Asn Asp Ser Gly Pro Thr Pro Gly Pro Ser	
275 280 285	
GCG CTC TTT GCC AGC CGC GTG CGC CCC GGG CAG CGC GTG TAC GTG GTG	912
Ala Leu Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val	
290 295 300	
GCT GAA CGC GGC GGG GAC CGC CGG CTG CTG CCC GCC GCG GTG CAC AGC	960
Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser	
305 310 315 320	
GTG ACG CTG CGA GAG GAG GAG GCG GGC GCG TAC GCG CCG CTC ACG GCG	1008
Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala	
325 330 335	
CAC GGC ACC ATT CTC ATC AAC CGG GTG CTC GCC TCG TGC TAC GCT GTC	1056
His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val	
340 345 350	
ATC GAG GAG CAC AGC TGG GCA CAC CGG GCC TTC GCG CCT TTC CGC CTG	1104
Ile Glu Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu	
355 360 365	
GCG CAC GCG CTG CTG GCC GCG CTG GCA CCC GCC CGC ACG GAC GGC GGG	1152
Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly	
370 375 380	
GGC GGG GGC AGC ATC CCT GCA GCG CAA TCT GCA ACG GAA GCG AGG GGC	1200
Gly Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly	
385 390 395 400	
GCG GAG CCG ACT GCG GGC ATC CAC TGG TAC TCG CAG CTG CTC TAC CAC	1248
Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His	
405 410 415	
ATT GGC ACC TGG CTG TTG GAC AGC GAG ACC ATG CAT CCC TTG GGA ATG	1296
Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met	
420 425 430	
GCG GTC AAG TCC AGC TG	1313
Ala Val Lys Ser Ser	
435	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1256 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS

E1
Cont

(B) LOCATION: 1..1254

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG CGG CTT TTG ACG AGA GTG CTG CTG GTG TCT CTT CTC ACT CTG TCC	48
Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser	
1 5 10 15	
TTG GTG GTG TCC GGA CTG GCC TGC GGT CCT GGC AGA GGC TAC GGC AGA	96
Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg	
20 25 30	
AGA AGA CAT CCG AAG AAG CTG ACA CCT CTC GCC TAC AAG CAG TTC ATA	144
Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile	
35 40 45	
CCT AAT GTC GCG GAG AAG ACC TTA GGG GCC AGC GGC AGA TAC GAG GGC	192
Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly	
50 55 60	
AAG ATA ACG CGC AAT TCG GAG AGA TTT AAA GAA CTT ACT CCA AAT TAC	240
Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr	
65 70 75 80	
AAT CCC GAC ATT ATC TTT AAG GAT GAG GAG AAC ACG GGA GCG GAC AGG	288
Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg	
85 90 95	
CTC ATG ACA CAG AGA TGC AAA GAC AAG CTG AAC TCG CTG GCC ATC TCT	336
Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser	
100 105 110	
GTA ATG AAC CAC TGG CCA GGG GTT AAG CTG CGT GTG ACA GAG GGC TGG	384
Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp	
115 120 125	
GAT GAG GAC GGT CAC CAT TTT GAA GAA TCA CTC CAC TAC GAG GGA AGA	432
Asp Glu Asp Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg	
130 135 140	
GCT GTT GAT ATT ACC ACC TCT GAC CGA GAC AAG AGC AAA TAC GGG ACA	480
Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr	
145 150 155 160	
CTG TCT CGC CTA GCT GTG GAG GCT GGA TTT GAC TGG GTC TAT TAC GAG	528
Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu	
165 170 175	
TCC AAA GCC CAC ATT CAT TGC TCT GTC AAA GCA GAA AAT TCG GTT GCT	576
Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala	
180 185 190	
GCG AAA TCT GGG GGC TGT TTC CCA GGT TCG GCT CTG GTC TCG CTC CAG	624
Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu Val Ser Leu Gln	
195 200 205	

E1
Cont.

GAC GGA GGA CAG AAG GCC GTG AAG GAC CTG AAC CCC GGA GAC AAG GTG Asp Gly Gly Gln Lys Ala Val Lys Asp Leu Asn Pro Gly Asp Lys Val 210 215 220	672
CTG GCG GCA GAC AGC GCG GGA AAC CTG GTG TTC AGC GAC TTC ATC ATG Leu Ala Ala Asp Ser Ala Gly Asn Leu Val Phe Ser Asp Phe Ile Met 225 230 235 240	720
TTC ACA GAC CGA GAC TCC ACG ACG CGA CGT GTG TTT TAC GTC ATA GAA Phe Thr Asp Arg Asp Ser Thr Thr Arg Arg Val Phe Tyr Val Ile Glu 245 250 255	768
ACG CAA GAA CCC GTT GAA AAG ATC ACC CTC ACC GCC GCT CAC CTC CTT Thr Gln Glu Pro Val Glu Lys Ile Thr Leu Thr Ala Ala His Leu Leu 260 265 270	816
TTT GTC CTC GAC AAC TCA ACG GAA GAT CTC CAC ACC ATG ACC GCC GCG Phe Val Leu Asp Asn Ser Thr Glu Asp Leu His Thr Met Thr Ala Ala 275 280 285	864
TAT GCC AGC AGT GTC AGA GCC GGA CAA AAG GTG ATG GTT GTT GAT GAT Tyr Ala Ser Ser Val Arg Ala Gly Gln Lys Val Met Val Val Asp Asp 290 295 300	912
AGC GGT CAG CTT AAA TCT GTC ATC GTG CAG CGG ATA TAC ACG GAG GAG Ser Gly Gln Leu Lys Ser Val Ile Val Gln Arg Ile Tyr Thr Glu Glu 305 310 315 320	960
CAG CGG GGC TCG TTC GCA CCA GTG ACT GCA CAT GGG ACC ATT GTG GTC Gln Arg Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile Val Val 325 330 335	1008
GAC AGA ATA CTG GCG TCC TGT TAC GCC GTA ATA GAG GAC CAG GGG CTT Asp Arg Ile Leu Ala Ser Cys Tyr Ala Val Ile Glu Asp Gln Gly Leu 340 345 350	1056
GCG CAT TTG GCC TTC GCG CCC GCC AGG CTC TAT TAT TAC GTG TCA TCA Ala His Leu Ala Phe Ala Pro Ala Arg Leu Tyr Tyr Tyr Val Ser Ser 355 360 365	1104
TTC CTG TCC CCC AAA ACT CCA GCA GTC GGT CCA ATG CGA CTT TAC AAC Phe Leu Ser Pro Lys Thr Pro Ala Val Gly Pro Met Arg Leu Tyr Asn 370 375 380	1152
AGG AGG GGG TCC ACT GGT ACT CCA GGC TCC TGT CAT CAA ATG GGA ACG Arg Arg Gly Ser Thr Gly Thr Pro Gly Ser Cys His Gln Met Gly Thr 385 390 395 400	1200
TGG CTT TTG GAC AGC AAC ATG CTT CAT CCT TTG GGG ATG TCA GTA AAC Trp Leu Leu Asp Ser Asn Met Leu His Pro Leu Gly Met Ser Val Asn 405 410 415	1248
TCA AGC TG Ser Ser	1256

(2) INFORMATION FOR SEQ ID NO:6:

E1
Cont.

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1425 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1425

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG CTG CTG CTG GCG AGA TGT CTG CTG CTA GTC CTC GTC TCC TCG CTG	48
Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu	
1 5 10 15	
CTG GTA TGC TCG GGA CTG GCG TGC GGA CCG GGC AGG GGG TTC GGG AAG	96
Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys	
20 25 30	
AGG AGG CAC CCC AAA AAG CTG ACC CCT TTA GCC TAC AAG CAG TTT ATC	144
Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile	
35 40 45	
CCC AAT GTG GCC GAG AAG ACC CTA GGC GCC AGC GGA AGG TAT GAA GGG	192
Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly	
50 55 60	
AAG ATC TCC AGA AAC TCC GAG CGA TTT AAG GAA CTC ACC CCC AAT TAC	240
Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr	
65 70 75 80	
AAC CCC GAC ATC ATA TTT AAG GAT GAA GAA AAC ACC GGA GCG GAC AGG	288
Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg	
85 90 95	
CTG ATG ACT CAG AGG TGT AAG GAC AAG TTG AAC GCT TTG GCC ATC TCG	336
Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser	
100 105 110	
GTG ATG AAC CAG TGG CCA GGA GTG AAA CTG CGG GTG ACC GAG GGC TGG	384
Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp	
115 120 125	
GAC GAA GAT GGC CAC CAC TCA GAG GAG TCT CTG CAC TAC GAG GGC CGC	432
Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg	
130 135 140	
GCA GTG GAC ATC ACC ACG TCT GAC CGC GAC CGC AGC AAG TAC GGC ATG	480
Ala Val Asp Ile Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met	
145 150 155 160	

E1
Cont

CTG	GCC	CGC	CTG	GCG	GTG	GAG	GCC	GGC	TTC	GAC	TGG	GTG	TAC	TAC	GAG	528
Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	Glu	
			165						170						175	
TCC	AAG	GCA	CAT	ATC	CAC	TGC	TCG	GTG	AAA	GCA	GAG	AAC	TCG	GTG	GCG	576
Ser	Lys	Ala	His	Ile	His	Cys	Ser	Val	Lys	Ala	Glu	Asn	Ser	Val	Ala	
			180					185					190			
GCC	AAA	TCG	GGA	GGC	TGC	TTC	CCG	GGC	TCG	GCC	ACG	GTG	CAC	CTG	GAG	624
Ala	Lys	Ser	Gly	Gly	Cys	Phe	Pro	Gly	Ser	Ala	Thr	Val	His	Leu	Glu	
		195					200					205				
CAG	GGC	GGC	ACC	AAG	CTG	GTG	AAG	GAC	CTG	AGC	CCC	GGG	GAC	CGC	GTG	672
Gln	Gly	Gly	Thr	Lys	Leu	Val	Lys	Asp	Leu	Ser	Pro	Gly	Asp	Arg	Val	
	210					215					220					
CTG	GCG	GCG	GAC	GAC	CAG	GGC	CGG	CTG	CTC	TAC	AGC	GAC	TTC	CTC	ACT	720
Leu	Ala	Ala	Asp	Asp	Gln	Gly	Arg	Leu	Leu	Tyr	Ser	Asp	Phe	Leu	Thr	
225					230					235					240	
TTC	CTG	GAC	CGC	GAC	GAC	GGC	GCC	AAG	AAG	GTC	TTC	TAC	GTG	ATC	GAG	768
Phe	Leu	Asp	Arg	Asp	Asp	Gly	Ala	Lys	Lys	Val	Phe	Tyr	Val	Ile	Glu	
			245					250						255		
ACG	CGG	GAG	CCG	CGC	GAG	CGC	CTG	CTG	CTC	ACC	GCC	GCG	CAC	CTG	CTC	816
Thr	Arg	Glu	Pro	Arg	Glu	Arg	Leu	Leu	Leu	Thr	Ala	Ala	His	Leu	Leu	
			260				265						270			
TTT	GTG	GCG	CCG	CAC	AAC	GAC	TCG	GCC	ACC	GGG	GAG	CCC	GAG	GCG	TCC	864
Phe	Val	Ala	Pro	His	Asn	Asp	Ser	Ala	Thr	Gly	Glu	Pro	Glu	Ala	Ser	
	275						280				285					
TCG	GGC	TCG	GGG	CCG	CCT	TCC	GGG	GGC	GCA	CTG	GGG	CCT	CGG	GCG	CTG	912
Ser	Gly	Ser	Gly	Pro	Pro	Ser	Gly	Gly	Ala	Leu	Gly	Pro	Arg	Ala	Leu	
	290					295					300					
TTC	GCC	AGC	CGC	GTG	CGC	CCG	GGC	CAG	CGC	GTG	TAC	GTG	GTG	GCC	GAG	960
Phe	Ala	Ser	Arg	Val	Arg	Pro	Gly	Gln	Arg	Val	Tyr	Val	Val	Ala	Glu	
305					310					315					320	
CGT	GAC	GGG	GAC	CGC	CGG	CTC	CTG	CCC	GCC	GCT	GTG	CAC	AGC	GTG	ACC	1008
Arg	Asp	Gly	Asp	Arg	Arg	Leu	Leu	Pro	Ala	Ala	Val	His	Ser	Val	Thr	
			325					330					335			
CTA	AGC	GAG	GAG	GCC	GCG	GGC	GCC	TAC	GCG	CCG	CTC	ACG	GCC	CAG	GGC	1056
Leu	Ser	Glu	Glu	Ala	Ala	Gly	Ala	Tyr	Ala	Pro	Leu	Thr	Ala	Gln	Gly	
			340				345						350			
ACC	ATT	CTC	ATC	AAC	CGG	GTG	CTG	GCC	TCG	TGC	TAC	GCG	GTC	ATC	GAG	1104
Thr	Ile	Leu	Ile	Asn	Arg	Val	Leu	Ala	Ser	Cys	Tyr	Ala	Val	Ile	Glu	
	355					360					365					
GAG	CAC	AGC	TGG	GCG	CAC	CGG	GCC	TTC	GCG	CCC	TTC	CGC	CTG	GCG	CAC	1152
Glu	His	Ser	Trp	Ala	His	Arg	Ala	Phe	Ala	Pro	Phe	Arg	Leu	Ala	His	
	370					375					380					

E1
Cont.

GCG CTC CTG GCT GCA CTG GCG CCC GCG CGC ACG GAC CGC GGC GGG GAC	1200
Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp	
385 390 395 400	
AGC GGC GGC GGG GAC CGC GGG GGC GGC GGC GGC AGA GTA GCC CTA ACC	1248
Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Arg Val Ala Leu Thr	
405 410 415	
GCT CCA GGT GCT GCC GAC GCT CCG GGT GCG GGG GCC ACC GCG GGC ATC	1296
Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile	
420 425 430	
CAC TGG TAC TCG CAG CTG CTC TAC CAA ATA GGC ACC TGG CTC CTG GAC	1344
His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp	
435 440 445	
AGC GAG GCC CTG CAC CCG CTG GGC ATG GCG GTC AAG TCC AGC NNN AGC	1392
Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser	
450 455 460	
CGG GGG GCC GGG GGA GGG GCG CGG GAG GGG GCC	1425
Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala	
465 470 475	

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 939 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..936

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGG CGC CTC ATG ACC CAG CGC TGC AAG GAC CGC CTG AAC TCG CTG GCT	48
Arg Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn Ser Leu Ala	
1 5 10 15	
ATC TCG GTG ATG AAC CAG TGG CCC GGT GTG AAG CTG CGG GTG ACC GAG	96
Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu	
20 25 30	
GGC TGG GAC GAG GAC GGC CAC CAC TCA GAG GAG TCC CTG CAT TAT GAG	144
Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu	
35 40 45	
GGC CGC GCG GTG GAC ATC ACC ACA TCA GAC CGC GAC CGC AAT AAG TAT	192
Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr	
50 55 60	

E1
Cont.

GGA	CTG	CTG	GCG	CGC	TTG	GCA	GTG	GAG	GCC	GGC	TTT	GAC	TGG	GTG	TAT	240
Gly	Leu	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	
65					70					75					80	
TAC	GAG	TCA	AAG	GCC	CAC	GTG	CAT	TGC	TCC	GTC	AAG	TCC	GAG	CAC	TCG	288
Tyr	Glu	Ser	Lys	Ala	His	Val	His	Cys	Ser	Val	Lys	Ser	Glu	His	Ser	
			85						90					95		
GCC	GCA	GCC	AAG	ACG	GGC	GGC	TGC	TTC	CCT	GCC	GGA	GCC	CAG	GTA	CGC	336
Ala	Ala	Ala	Lys	Thr	Gly	Gly	Cys	Phe	Pro	Ala	Gly	Ala	Gln	Val	Arg	
			100					105					110			
CTG	GAG	AGT	GGG	GCG	CGT	GTG	GCC	TTG	TCA	GCC	GTG	AGG	CCG	GGA	GAC	384
Leu	Glu	Ser	Gly	Ala	Arg	Val	Ala	Leu	Ser	Ala	Val	Arg	Pro	Gly	Asp	
		115					120					125				
CGT	GTG	CTG	GCC	ATG	GGG	GAG	GAT	GGG	AGC	CCC	ACC	TTC	AGC	GAT	GTG	432
Arg	Val	Leu	Ala	Met	Gly	Glu	Asp	Gly	Ser	Pro	Thr	Phe	Ser	Asp	Val	
	130					135					140					
CTC	ATT	TTC	CTG	GAC	CGC	GAG	CCC	CAC	AGG	CTG	AGA	GCC	TTC	CAG	GTC	480
Leu	Ile	Phe	Leu	Asp	Arg	Glu	Pro	His	Arg	Leu	Arg	Ala	Phe	Gln	Val	
145					150				155						160	
ATC	GAG	ACT	CAG	GAC	CCC	CCA	CGC	CGC	CTG	GCA	CTC	ACA	CCC	GCT	CAC	528
Ile	Glu	Thr	Gln	Asp	Pro	Pro	Arg	Arg	Leu	Ala	Leu	Thr	Pro	Ala	His	
			165						170					175		
CTG	CTC	TTT	ACG	GCT	GAC	AAT	CAC	ACG	GAG	CCG	GCA	GCC	CGC	TTC	CGG	576
Leu	Leu	Phe	Thr	Ala	Asp	Asn	His	Thr	Glu	Pro	Ala	Ala	Arg	Phe	Arg	
			180					185					190			
GCC	ACA	TTT	GCC	AGC	CAC	GTG	CAG	CCT	GGC	CAG	TAC	GTG	CTG	GTG	GCT	624
Ala	Thr	Phe	Ala	Ser	His	Val	Gln	Pro	Gly	Gln	Tyr	Val	Leu	Val	Ala	
		195					200					205				
GGG	GTG	CCA	GGC	CTG	CAG	CCT	GCC	CGC	GTG	GCA	GCT	GTC	TCT	ACA	CAC	672
Gly	Val	Pro	Gly	Leu	Gln	Pro	Ala	Arg	Val	Ala	Ala	Val	Ser	Thr	His	
	210					215					220					
GTG	GCC	CTC	GGG	GCC	TAC	GCC	CCG	CTC	ACA	AAG	CAT	GGG	ACA	CTG	GTG	720
Val	Ala	Leu	Gly	Ala	Tyr	Ala	Pro	Leu	Thr	Lys	His	Gly	Thr	Leu	Val	
225					230					235					240	
GTG	GAG	GAT	GTG	GTG	GCA	TCC	TGC	TTC	GCG	GCC	GTG	GCT	GAC	CAC	CAC	768
Val	Glu	Asp	Val	Val	Ala	Ser	Cys	Phe	Ala	Ala	Val	Ala	Asp	His	His	
			245						250					255		
CTG	GCT	CAG	TTG	GCC	TTC	TGG	CCC	CTG	AGA	CTC	TTT	CAC	AGC	TTG	GCA	816
Leu	Ala	Gln	Leu	Ala	Phe	Trp	Pro	Leu	Arg	Leu	Phe	His	Ser	Leu	Ala	
			260					265					270			
TGG	GGC	AGC	TGG	ACC	CCG	GGG	GAG	GGT	GTG	CAT	TGG	TAC	CCC	CAG	CTG	864
Trp	Gly	Ser	Trp	Thr	Pro	Gly	Glu	Gly	Val	His	Trp	Tyr	Pro	Gln	Leu	
		275					280					285				

E1
Cont

CTC TAC CGC CTG GGG CGT CTC CTG CTA GAA GAG GGC AGC TTC CAC CCA
 Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Gly Ser Phe His Pro
 290 295 300

912

CTG GGC ATG TCC GGG GCA GGG AGC TGA
 Leu Gly Met Ser Gly Ala Gly Ser
 305 310

939

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 425 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile
 1 5 10 15
 Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly
 20 25 30
 Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
 35 40 45
 Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg
 50 55 60
 Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr
 65 70 75 80
 Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly
 85 90 95
 Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu
 100 105 110
 Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr
 115 120 125
 Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr
 130 135 140
 Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys
 145 150 155 160
 Tyr Gly Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val
 165 170 175
 Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn
 180 185 190
 Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val
 195 200 205

E1
cont

His Leu Glu His Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly
 210 215 220
 Asp Arg Val Leu Ala Ala Asp Ala Asp Gly Arg Leu Leu Tyr Ser Asp
 225 230 235 240
 Phe Leu Thr Phe Leu Asp Arg Met Asp Ser Ser Arg Lys Leu Phe Tyr
 245 250 255
 Val Ile Glu Thr Arg Gln Pro Arg Ala Arg Leu Leu Leu Thr Ala Ala
 260 265 270
 His Leu Leu Phe Val Ala Pro Gln His Asn Gln Ser Glu Ala Thr Gly
 275 280 285
 Ser Thr Ser Gly Gln Ala Leu Phe Ala Ser Asn Val Lys Pro Gly Gln
 290 295 300
 Arg Val Tyr Val Leu Gly Glu Gly Gly Gln Gln Leu Leu Pro Ala Ser
 305 310 315 320
 Val His Ser Val Ser Leu Arg Glu Glu Ala Ser Gly Ala Tyr Ala Pro
 325 330 335
 Leu Thr Ala Gln Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys
 340 345 350
 Tyr Ala Val Ile Glu Glu His Ser Trp Ala His Trp Ala Phe Ala Pro
 355 360 365
 Phe Arg Leu Ala Gln Gly Leu Leu Ala Ala Leu Cys Pro Asp Gly Ala
 370 375 380
 Ile Pro Thr Ala Ala Thr Thr Thr Thr Gly Ile His Trp Tyr Ser Arg
 385 390 395 400
 Leu Leu Tyr Arg Ile Gly Ser Trp Val Leu Asp Gly Asp Ala Leu His
 405 410 415
 Pro Leu Gly Met Val Ala Pro Ala Ser
 420 425

E1
Cont

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 396 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Ala Leu Pro Ala Ser Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu
 1 5 10 15

Ala Leu Ser Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg
 20 25 30
 Arg Arg Tyr Val Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe
 35 40 45
 Val Pro Ser Met Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu
 50 55 60
 Gly Arg Val Thr Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn
 65 70 75 80
 Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp
 85 90 95
 Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile
 100 105 110
 Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
 115 120 125
 Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly
 130 135 140
 Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly
 145 150 155 160
 Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
 165 170 175
 Glu Ser Arg Asn His Ile His Val Ser Val Lys Ala Asp Asn Ser Leu
 180 185 190
 Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu
 195 200 205
 Arg Ser Gly Glu Arg Lys Gly Leu Arg Glu Leu His Arg Gly Asp Trp
 210 215 220
 Val Leu Ala Ala Asp Ala Ala Gly Arg Val Val Pro Thr Pro Val Leu
 225 230 235 240
 Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val
 245 250 255
 Glu Thr Glu Arg Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu
 260 265 270
 Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro
 275 280 285
 Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly
 290 295 300
 Gly Asp Ala Leu Gln Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu
 305 310 315 320

E
 Cont

Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
 325 330 335

Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
 340 345 350

Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala
 355 360 365

Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser
 370 375 380

Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Met Gly
 385 390 395

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu
 1 5 10 15

Leu Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg
 20 25 30

Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala
 35 40 45

Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser
 50 55 60

Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu
 65 70 75 80

Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn
 85 90 95

Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn
 100 105 110

Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg
 115 120 125

Val Thr Glu Gly Arg Asp Glu Asp Gly His His Ser Glu Glu Ser Leu
 130 135 140

His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg
 145 150 155 160

E1
Cont

Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp
 165 170 175
 Trp Val Tyr Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser
 180 185 190
 Glu His Ser Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala
 195 200 205
 Gln Val Arg Leu Glu Asn Gly Glu Arg Val Ala Leu Ser Ala Val Lys
 210 215 220
 Pro Gly Asp Arg Val Leu Ala Met Gly Glu Asp Gly Thr Pro Thr Phe
 225 230 235 240
 Ser Asp Val Leu Ile Phe Leu Asp Arg Glu Pro Asn Arg Leu Arg Ala
 245 250 255
 Phe Gln Val Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr
 260 265 270
 Pro Ala His Leu Leu Phe Ile Ala Asp Asn His Thr Glu Pro Ala Ala
 275 280 285
 His Phe Arg Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val
 290 295 300
 Leu Val Ser Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val
 305 310 315 320
 Ser Thr His Val Ala Leu Gly Ser Tyr Ala Pro Leu Thr Arg His Gly
 325 330 335
 Thr Leu Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala
 340 345 350
 Asp His His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe Pro
 355 360 365
 Ser Leu Ala Trp Gly Ser Trp Thr Pro Ser Glu Gly Val His Ser Tyr
 370 375 380
 Pro Gln Met Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Ser Thr
 385 390 395 400
 Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser
 405 410

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 437 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

E1
cont.

Met Leu Leu Leu Leu Ala Arg Cys Phe Leu Val Ile Leu Ala Ser Ser
 1 5 10 15
 Leu Leu Val Cys Pro Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly
 20 25 30
 Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe
 35 40 45
 Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu
 50 55 60
 Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn
 65 70 75 80
 Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp
 85 90 95
 Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile
 100 105 110
 Ser Val Met Asn Gln Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
 115 120 125
 Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly
 130 135 140
 Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly
 145 150 155 160
 Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
 165 170 175
 Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val
 180 185 190
 Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu
 195 200 205
 Glu Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Arg Pro Gly Asp Arg
 210 215 220
 Val Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu
 225 230 235 240
 Thr Phe Leu Asp Arg Asp Glu Gly Ala Lys Lys Val Phe Tyr Val Ile
 245 250 255
 Glu Thr Leu Glu Pro Arg Glu Arg Leu Leu Thr Ala Ala His Leu
 260 265 270
 Leu Phe Val Ala Pro His Asn Asp Ser Gly Pro Thr Pro Gly Pro Ser
 275 280 285
 Ala Leu Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val
 290 295 300

E!
 Cont.

Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser
 305 310 315 320

Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala
 325 330 335

His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val
 340 345 350

Ile Glu Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu
 355 360 365

Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly
 370 375 380

Gly Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly
 385 390 395 400

Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His
 405 410 415

Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met
 420 425 430

Ala Val Lys Ser Ser
 435

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 418 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser
 1 5 10 15

Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg
 20 25 30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
 35 40 45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
 50 55 60

Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
 65 70 75 80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
 85 90 95

E1
Cont

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser
 100 105 110
 Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
 115 120 125
 Asp Glu Asp Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg
 130 135 140
 Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr
 145 150 155 160
 Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
 165 170 175
 Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
 180 185 190
 Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu Val Ser Leu Gln
 195 200 205
 Asp Gly Gly Gln Lys Ala Val Lys Asp Leu Asn Pro Gly Asp Lys Val
 210 215 220
 Leu Ala Ala Asp Ser Ala Gly Asn Leu Val Phe Ser Asp Phe Ile Met
 225 230 235 240
 Phe Thr Asp Arg Asp Ser Thr Thr Arg Arg Val Phe Tyr Val Ile Glu
 245 250 255
 Thr Gln Glu Pro Val Glu Lys Ile Thr Leu Thr Ala Ala His Leu Leu
 260 265 270
 Phe Val Leu Asp Asn Ser Thr Glu Asp Leu His Thr Met Thr Ala Ala
 275 280 285
 Tyr Ala Ser Ser Val Arg Ala Gly Gln Lys Val Met Val Val Asp Asp
 290 295 300
 Ser Gly Gln Leu Lys Ser Val Ile Val Gln Arg Ile Tyr Thr Glu Glu
 305 310 315 320
 Gln Arg Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile Val Val
 325 330 335
 Asp Arg Ile Leu Ala Ser Cys Tyr Ala Val Ile Glu Asp Gln Gly Leu
 340 345 350
 Ala His Leu Ala Phe Ala Pro Ala Arg Leu Tyr Tyr Tyr Val Ser Ser
 355 360 365
 Phe Leu Ser Pro Lys Thr Pro Ala Val Gly Pro Met Arg Leu Tyr Asn
 370 375 380
 Arg Arg Gly Ser Thr Gly Thr Pro Gly Ser Cys His Gln Met Gly Thr
 385 390 395 400

E1
 Cont

Trp Leu Leu Asp Ser Asn Met Leu His Pro Leu Gly Met Ser Val Asn
 405 410 415

Ser Ser

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 475 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu
 1 5 10 15

Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys
 20 25 30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
 35 40 45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
 50 55 60

Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
 65 70 75 80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
 85 90 95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser
 100 105 110

Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
 115 120 125

Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg
 130 135 140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met
 145 150 155 160

Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
 165 170 175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
 180 185 190

Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu
 195 200 205

Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val
 210 215 220

E1
 Cont

Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr
 225 230 235 240
 Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu
 245 250 255
 Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu
 260 265 270
 Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser
 275 280 285
 Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu
 290 295 300
 Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu
 305 310 315 320
 Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr
 325 330 335
 Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly
 340 345 350
 Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu
 355 360 365
 Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His
 370 375 380
 Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp
 385 390 395 400
 Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Arg Val Ala Leu Thr
 405 410 415
 Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile
 420 425 430
 His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp
 435 440 445
 Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser
 450 455 460
 Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala
 465 470 475

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

E1
Cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Arg Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn Ser Leu Ala
 1 5 10 15
 Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu
 20 25 30
 Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu
 35 40 45
 Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr
 50 55 60
 Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr
 65 70 75 80
 Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser Glu His Ser
 85 90 95
 Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala Gln Val Arg
 100 105 110
 Leu Glu Ser Gly Ala Arg Val Ala Leu Ser Ala Val Arg Pro Gly Asp
 115 120 125
 Arg Val Leu Ala Met Gly Glu Asp Gly Ser Pro Thr Phe Ser Asp Val
 130 135 140
 Leu Ile Phe Leu Asp Arg Glu Pro His Arg Leu Arg Ala Phe Gln Val
 145 150 155 160
 Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr Pro Ala His
 165 170 175
 Leu Leu Phe Thr Ala Asp Asn His Thr Glu Pro Ala Ala Arg Phe Arg
 180 185 190
 Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val Leu Val Ala
 195 200 205
 Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val Ser Thr His
 210 215 220
 Val Ala Leu Gly Ala Tyr Ala Pro Leu Thr Lys His Gly Thr Leu Val
 225 230 235 240
 Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala Asp His His
 245 250 255
 Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe His Ser Leu Ala
 260 265 270
 Trp Gly Ser Trp Thr Pro Gly Glu Gly Val His Trp Tyr Pro Gln Leu
 275 280 285

E1
 Cont

Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Gly Ser Phe His Pro
 290 295 300

Leu Gly Met Ser Gly Ala Gly Ser
 305 310

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 64 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser Val Met Asn
 1 5 10 15

His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
 20 25 30

Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp
 35 40 45

Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr Leu Ser Arg
 50 55 60

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 65 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gln Arg Cys Lys Glu Lys Leu Asn Ser Leu Ala Ile Ser Val Met Asn
 1 5 10 15

Met Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
 20 25 30

*F1
Cont*

Gly Asn His Phe Glu Asp Ser Leu His Tyr Glu Gly Arg Ala Val Asp
 35 40 45

Ile Thr Thr Ser Ser Asp Arg Asp Arg Asn Lys Tyr Gly Met Phe Ala
 50 55 60

Arg
 65

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 64 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser Val Met Asn
 1 5 10 15

Leu Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
 20 25 30

Gly Leu His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp
 35 40 45

Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Arg Met Leu Ala Arg
 50 55 60

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGAATTCCCA GMGNTGYAAR GARMRNBNA

30

(2) INFORMATION FOR SEQ ID NO:19:

E1
 Cont.

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCATCGATGG ACCCARTCRA ANCCNGCYTC

30

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCTCTAGAGC TCNACNGCNA RNCKNGC

27

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AGCTGTGCGAC GCGGCCGCTA CGTAGGTTAC CGACGTCAAG CTTAGATCTC

50

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

E1
Cont

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCTGAGATC TAAGCTTGAC GTCGGTAACC TACGTAGCGG CCGCGTCGAC

50

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCGGCCAG GCAGGCCTCG CGATATCGTC ACCGCGGTAT TCGAA

45

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGTGCCAGTC GGGGCCCCCA GGGCCGCGCC

30

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

E1
Cont

TACCACAGCG GATGGTTCGG

20

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTGGTGGTTA TGCCGATCGC

20

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TAAGAGGCCT ATAAGAGGCG G

21

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAGTCAGCCC AGAGGAGACT

20

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid

E1
Cont

- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Cys Gly Pro Gly Arg Gly
1 5

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AGMGNTGYAA RGARMRNBNA A

21

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CTCNACNGCN ARNCKNGCNA

20

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

E1
Cont

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTGCAGGGAT CCACCATGCG GCTTTTGACG AG

32

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTGCAGGGAT CCTTATTCCA CACGAGGGAT T

31

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Asp Asn His Ser Ser Val Pro Trp Ala Ser Ala Ala Ser Val Thr
1 5 10 15

Cys Leu Ser Leu Asp Ala Lys Cys His Ser Ser Ser Ser Ser Ser Ser
20 25 30

Ser Lys Ser Ala Ala Ser Ser Ile Ser Ala Ile Pro Gln Glu Glu Thr
35 40 45

Gln Thr Met Arg His Ile Ala His Thr Gln Arg Cys Leu Ser Arg Leu
50 55 60

Thr Ser Leu Val Ala Leu Leu Leu Ile Val Leu Pro Met Val Phe Ser
65 70 75 80

E1
Cont.

Pro Ala His Ser Cys Gly Pro Gly Arg Gly Leu Gly Arg His Arg Ala
 85 90 95
 Arg Asn Leu Tyr Pro Leu Val Leu Lys Gln Thr Ile Pro Asn Leu Ser
 100 105 110
 Glu Tyr Thr Asn Ser Ala Ser Gly Pro Leu Glu Gly Val Ile Arg Arg
 115 120 125
 Asp Ser Pro Lys Phe Lys Asp Leu Val Pro Asn Tyr Asn Arg Asp Ile
 130 135 140
 Leu Phe Arg Asp Glu Glu Gly Thr Gly Ala Asp Arg Leu Met Ser Lys
 145 150 155 160
 Arg Cys Lys Glu Lys Leu Asn Val Leu Ala Tyr Ser Val Met Asn Glu
 165 170 175
 Trp Pro Gly Ile Arg Leu Leu Val Thr Glu Ser Trp Asp Glu Asp Tyr
 180 185 190
 His His Gly Gln Glu Ser Leu His Tyr Glu Gly Arg Ala Val Thr Ile
 195 200 205
 Ala Thr Ser Asp Arg Asp Gln Ser Lys Tyr Gly Met Leu Ala Arg Leu
 210 215 220
 Ala Val Glu Ala Gly Phe Asp Trp Val Ser Tyr Val Ser Arg Arg His
 225 230 235 240
 Ile Tyr Cys Ser Val Lys Ser Asp Ser Ser Ile Ser Ser His Val His
 245 250 255
 Gly Cys Phe Thr Pro Glu Ser Thr Ala Leu Leu Glu Ser Gly Val Arg
 260 265 270
 Lys Pro Leu Gly Glu Leu Ser Ile Gly Asp Arg Val Leu Ser Met Thr
 275 280 285
 Ala Asn Gly Gln Ala Val Tyr Ser Glu Val Ile Leu Phe Met Asp Arg
 290 295 300
 Asn Leu Glu Gln Met Gln Asn Phe Val Gln Leu His Thr Asp Gly Gly
 305 310 315 320
 Ala Val Leu Thr Val Thr Pro Ala His Leu Val Ser Val Trp Gln Pro
 325 330 335
 Glu Ser Gln Lys Leu Thr Phe Val Phe Ala Asp Arg Ile Glu Glu Lys
 340 345 350
 Asn Gln Val Leu Val Arg Asp Val Glu Thr Gly Glu Leu Arg Pro Gln
 355 360 365
 Arg Val Val Lys Val Gly Ser Val Arg Ser Lys Gly Val Val Ala Pro
 370 375 380

E1
 Cont.

Leu Thr Arg Glu Gly Thr Ile Val Val Asn Ser Val Ala Ala Ser Cys
 385 390 395 400
 Tyr Ala Val Ile Asn Ser Gln Ser Leu Ala His Trp Gly Leu Ala Pro
 405 410 415
 Met Arg Leu Leu Ser Thr Leu Glu Ala Trp Leu Pro Ala Lys Glu Gln
 420 425 430
 Leu His Ser Ser Pro Lys Val Val Ser Ser Ala Gln Gln Gln Asn Gly
 435 440 445
 Ile His Trp Tyr Ala Asn Ala Leu Tyr Lys Val Lys Asp Tyr Val Leu
 450 455 460
 Pro Gln Ser Trp Arg His Asp
 465 470

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 73 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Arg Cys Lys Glu Arg Val Asn Ser Leu Ala Ile Ala Val Met His Met
 1 5 10 15
 Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly
 20 25 30
 His His Leu Pro Asp Ser Leu His Tyr Glu Gly Arg Ala Leu Asp Ile
 35 40 45
 Thr Thr Ser Asp Arg Asp Arg His Lys Tyr Gly Met Leu Ala Arg Leu
 50 55 60
 Ala Val Glu Ala Gly Phe Asp Trp Val
 65 70

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 73 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

E1
Cont

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Arg	Cys	Lys	Asp	Lys	Leu	Asn	Ala	Leu	Ala	Ile	Ser	Val	Met	Asn	Gln
1				5					10					15	
Trp	Pro	Gly	Val	Lys	Leu	Arg	Val	Thr	Glu	Gly	Trp	Asp	Glu	Asp	Gly
			20					25					30		
His	His	Ser	Glu	Glu	Ser	Leu	His	Tyr	Glu	Gly	Arg	Ala	Val	Asp	Ile
			35				40					45			
Thr	Thr	Ser	Asp	Arg	Asp	Arg	Ser	Lys	Tyr	Gly	Met	Leu	Ala	Arg	Leu
			50			55					60				
Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val							
65						70									

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 64 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Lys	Arg	Cys	Lys	Glu	Lys	Leu	Asn	Val	Leu	Ala	Tyr	Ser	Val	Met	Asn
1				5					10					15	
Glu	Trp	Pro	Gly	Ile	Arg	Leu	Val	Val	Thr	Glu	Ser	Trp	Asp	Glu	Asp
			20					25					30		
Tyr	His	His	Gly	Gln	Glu	Ser	Leu	His	Tyr	Glu	Gly	Arg	Ala	Val	Thr
			35				40					45			
Ile	Ala	Thr	Ser	Asp	Arg	Asp	Gln	Ser	Lys	Tyr	Gly	Met	Leu	Ala	Arg
			50			55					60				

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs

E1
cont

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AAAAGCTTTA YTGTYAYGTN GGNATHGG

28

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AAGAATTCTA NGCRTTRTAR TTRTTNGG

28

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 221 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Cys Gly Pro Gly Arg Gly Xaa Gly Xaa Arg Arg His Pro Lys Lys Leu
1 5 10 15

Thr Pro Leu Ala Tyr Lys Gln Phe Ile Pro Asn Val Ala Glu Lys Thr
20 25 30

Leu Gly Ala Ser Gly Arg Tyr Glu Gly Lys Ile Xaa Arg Asn Ser Glu
35 40 45

Arg Phe Lys Glu Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys
50 55 60

E1
Cont.

Asp	Glu	Glu	Asn	Thr	Gly	Ala	Asp	Arg	Leu	Met	Thr	Gln	Arg	Cys	Lys
65					70					75					80
Asp	Lys	Leu	Asn	Xaa	Leu	Ala	Ile	Ser	Val	Met	Asn	Xaa	Trp	Pro	Gly
			85						90					95	
Val	Xaa	Leu	Arg	Val	Thr	Glu	Gly	Trp	Asp	Glu	Asp	Gly	His	His	Xaa
			100					105					110		
Glu	Glu	Ser	Leu	His	Tyr	Glu	Gly	Arg	Ala	Val	Asp	Ile	Thr	Thr	Ser
		115					120					125			
Asp	Arg	Asp	Xaa	Ser	Lys	Tyr	Gly	Xaa	Leu	Xaa	Arg	Leu	Ala	Val	Glu
	130					135					140				
Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	Glu	Ser	Lys	Ala	His	Ile	His	Cys
145					150					155					160
Ser	Val	Lys	Ala	Glu	Asn	Ser	Val	Ala	Ala	Lys	Ser	Gly	Gly	Cys	Phe
				165					170					175	
Pro	Gly	Ser	Ala	Xaa	Val	Xaa	Leu	Xaa	Xaa	Gly	Gly	Xaa	Lys	Xaa	Val
			180					185					190		
Lys	Asp	Leu	Xaa	Pro	Gly	Asp	Xaa	Val	Leu	Ala	Ala	Asp	Xaa	Xaa	Gly
		195					200					205			
Xaa	Leu	Xaa	Xaa	Ser	Asp	Phe	Xaa	Xaa	Phe	Xaa	Asp	Arg			
	210					215					220				

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 167 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Cys	Gly	Pro	Gly	Arg	Gly	Xaa	Xaa	Xaa	Arg	Arg	Xaa	Xaa	Xaa	Pro	Lys
1				5					10					15	
Xaa	Leu	Xaa	Pro	Leu	Xaa	Tyr	Lys	Gln	Phe	Xaa	Pro	Xaa	Xaa	Xaa	Glu
			20					25					30		
Xaa	Thr	Leu	Gly	Ala	Ser	Gly	Xaa	Xaa	Glu	Gly	Xaa	Xaa	Xaa	Arg	Xaa
		35				40					45				
Ser	Glu	Arg	Phe	Xaa	Xaa	Leu	Thr	Pro	Asn	Tyr	Asn	Pro	Asp	Ile	Ile
	50					55					60				

El
Cont.

Phe Lys Asp Glu Glu Asn Xaa Gly Ala Asp Arg Leu Met Thr Xaa Arg
 65 70 75 80
 Cys Lys Xaa Xaa Xaa Asn Xaa Leu Ala Ile Ser Val Met Asn Xaa Trp
 85 90 95
 Pro Gly Val Xaa Leu Arg Val Thr Glu Gly Xaa Asp Glu Asp Gly His
 100 105 110
 His Xaa Xaa Xaa Ser Leu His Tyr Glu Gly Arg Ala Xaa Asp Ile Thr
 115 120 125
 Thr Ser Asp Arg Asp Xaa Xaa Lys Tyr Gly Xaa Leu Xaa Arg Leu Ala
 130 135 140
 Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Xaa Xaa His Xaa
 145 150 155 160
 His Xaa Ser Val Lys Xaa Xaa
 165

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3900 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..3897

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATG GAC CGC GAC AGC CTC CCA CGC GTT CCG GAC ACA CAC GGC GAT GTG 48
 Met Asp Arg Asp Ser Leu Pro Arg Val Pro Asp Thr His Gly Asp Val
 1 5 10 15
 GTC GAT GAG AAA TTA TTC TCG GAT CTT TAC ATA CGC ACC AGC TGG GTG 96
 Val Asp Glu Lys Leu Phe Ser Asp Leu Tyr Ile Arg Thr Ser Trp Val
 20 25 30
 GAC GCC CAA GTG GCG CTC GAT CAG ATA GAT AAG GGC AAA GCG CGT GGC 144
 Asp Ala Gln Val Ala Leu Asp Gln Ile Asp Lys Gly Lys Ala Arg Gly
 35 40 45
 AGC CGC ACG GCG ATC TAT CTG CGA TCA GTA TTC CAG TCC CAC CTC GAA 192
 Ser Arg Thr Ala Ile Tyr Leu Arg Ser Val Phe Gln Ser His Leu Glu
 50 55 60

E1
Cont.

ACC	CTC	GGC	AGC	TCC	GTG	CAA	AAG	CAC	GCG	GGC	AAG	GTG	CTA	TTC	GTG	240
Thr	Leu	Gly	Ser	Ser	Val	Gln	Lys	His	Ala	Gly	Lys	Val	Leu	Phe	Val	
65					70					75					80	
GCT	ATC	CTG	GTG	CTG	AGC	ACC	TTC	TGC	GTC	GGC	CTG	AAG	AGC	GCC	CAG	288
Ala	Ile	Leu	Val	Leu	Ser	Thr	Phe	Cys	Val	Gly	Leu	Lys	Ser	Ala	Gln	
			85						90					95		
ATC	CAC	TCC	AAG	GTG	CAC	CAG	CTG	TGG	ATC	CAG	GAG	GGC	GGC	GGG	CTG	336
Ile	His	Ser	Lys	Val	His	Gln	Leu	Trp	Ile	Gln	Glu	Gly	Gly	Gly	Leu	
			100					105						110		
GAG	GCG	GAA	CTG	GCC	TAC	ACA	CAG	AAG	ACG	ATC	GGC	GAG	GAC	GAG	TCG	384
Glu	Ala	Glu	Leu	Ala	Tyr	Thr	Gln	Lys	Thr	Ile	Gly	Glu	Asp	Glu	Ser	
		115					120					125				
GCC	ACG	CAT	CAG	CTG	CTC	ATT	CAG	ACG	ACC	CAC	GAC	CCG	AAC	GCC	TCC	432
Ala	Thr	His	Gln	Leu	Leu	Ile	Gln	Thr	Thr	His	Asp	Pro	Asn	Ala	Ser	
		130				135					140					
GTC	CTG	CAT	CCG	CAG	GCG	CTG	CTT	GCC	CAC	CTG	GAG	GTC	CTG	GTC	AAG	480
Val	Leu	His	Pro	Gln	Ala	Leu	Leu	Ala	His	Leu	Glu	Val	Leu	Val	Lys	
145					150					155					160	
GCC	ACC	GCC	GTC	AAG	GTG	CAC	CTC	TAC	GAC	ACC	GAA	TGG	GGG	CTG	CGC	528
Ala	Thr	Ala	Val	Lys	Val	His	Leu	Tyr	Asp	Thr	Glu	Trp	Gly	Leu	Arg	
				165					170					175		
GAC	ATG	TGC	AAC	ATG	CCG	AGC	ACG	CCC	TCC	TTC	GAG	GGC	ATC	TAC	TAC	576
Asp	Met	Cys	Asn	Met	Pro	Ser	Thr	Pro	Ser	Phe	Glu	Gly	Ile	Tyr	Tyr	
			180					185					190			
ATC	GAG	CAG	ATC	CTG	CGC	CAC	CTC	ATT	CCG	TGC	TCG	ATC	ATC	ACG	CCG	624
Ile	Glu	Gln	Ile	Leu	Arg	His	Leu	Ile	Pro	Cys	Ser	Ile	Ile	Thr	Pro	
		195					200					205				
CTG	GAC	TGT	TTC	TGG	GAG	GGA	AGC	CAG	CTG	TTG	GGT	CCG	GAA	TCA	GCG	672
Leu	Asp	Cys	Phe	Trp	Glu	Gly	Ser	Gln	Leu	Leu	Gly	Pro	Glu	Ser	Ala	
	210					215					220					
GTC	GTT	ATA	CCA	GGC	CTC	AAC	CAA	CGA	CTC	CTG	TGG	ACC	ACA	CTG	AAT	720
Val	Val	Ile	Pro	Gly	Leu	Asn	Gln	Arg	Leu	Leu	Trp	Thr	Thr	Leu	Asn	
225					230					235					240	
CCC	GCC	TCT	GTG	ATG	CAG	TAT	ATG	AAG	CAG	AAG	ATG	TCC	GAG	GAA	AAG	768
Pro	Ala	Ser	Val	Met	Gln	Tyr	Met	Lys	Gln	Lys	Met	Ser	Glu	Glu	Lys	
				245				250						255		
ATC	AGC	TTC	GAC	TTC	GAG	ACC	GTG	GAG	CAG	TAC	ATG	AAG	CGT	GCG	GCC	816
Ile	Ser	Phe	Asp	Phe	Glu	Thr	Val	Glu	Gln	Tyr	Met	Lys	Arg	Ala	Ala	
			260					265					270			
ATT	GCG	AGT	GGC	TAC	ATG	GAG	AAG	CCC	TGC	CTG	AAC	CCA	CTG	AAT	CCC	864
Ile	Ala	Ser	Gly	Tyr	Met	Glu	Lys	Pro	Cys	Leu	Asn	Pro	Leu	Asn	Pro	
		275					280					285				

E1
Cont

AAT	TGC	CCG	GAC	ACG	GCA	CCG	AAC	AAG	AAC	AGC	ACC	CAG	CCG	CCG	GAT	912
Asn	Cys	Pro	Asp	Thr	Ala	Pro	Asn	Lys	Asn	Ser	Thr	Gln	Pro	Pro	Asp	
290						295					300					
GTG	GGA	GCC	ATC	CTG	TCC	GGA	GGC	TGC	TAC	GGT	TAT	GCC	GCG	AAG	CAC	960
Val	Gly	Ala	Ile	Leu	Ser	Gly	Gly	Cys	Tyr	Gly	Tyr	Ala	Ala	Lys	His	
305					310					315					320	
ATG	CAC	TGG	CCG	GAG	GAG	CTG	ATT	GTG	GGC	GGA	GCG	AAG	AGG	AAC	CGC	1008
Met	His	Trp	Pro	Glu	Glu	Leu	Ile	Val	Gly	Gly	Ala	Lys	Arg	Asn	Arg	
				325				330						335		
AGC	GGA	CAC	TTG	AGG	AAG	GCC	CAG	GCC	CTG	CAG	TCG	GTG	GTG	CAG	CTG	1056
Ser	Gly	His	Leu	Arg	Lys	Ala	Gln	Ala	Leu	Gln	Ser	Val	Val	Gln	Leu	
			340					345						350		
ATG	ACC	GAG	AAG	GAA	ATG	TAC	GAC	CAG	TGG	CAG	GAC	AAC	TAC	AAG	GTG	1104
Met	Thr	Glu	Lys	Glu	Met	Tyr	Asp	Gln	Trp	Gln	Asp	Asn	Tyr	Lys	Val	
		355					360					365				
CAC	CAT	CTT	GGA	TGG	ACG	CAG	GAG	AAG	GCA	GCG	GAG	GTT	TTG	AAC	GCC	1152
His	His	Leu	Gly	Trp	Thr	Gln	Glu	Lys	Ala	Ala	Glu	Val	Leu	Asn	Ala	
		370				375					380					
TGG	CAG	CGC	AAC	TTT	TCG	CGG	GAG	GTG	GAA	CAG	CTG	CTA	CGT	AAA	CAG	1200
Trp	Gln	Arg	Asn	Phe	Ser	Arg	Glu	Val	Glu	Gln	Leu	Leu	Arg	Lys	Gln	
385					390					395					400	
TCG	AGA	ATT	GCC	ACC	AAC	TAC	GAT	ATC	TAC	GTG	TTC	AGC	TCG	GCT	GCA	1248
Ser	Arg	Ile	Ala	Thr	Asn	Tyr	Asp	Ile	Tyr	Val	Phe	Ser	Ser	Ala	Ala	
				405					410					415		
CTG	GAT	GAC	ATC	CTG	GCC	AAG	TTC	TCC	CAT	CCC	AGC	GCC	TTG	TCC	ATT	1296
Leu	Asp	Asp	Ile	Leu	Ala	Lys	Phe	Ser	His	Pro	Ser	Ala	Leu	Ser	Ile	
			420					425						430		
GTC	ATC	GGC	GTG	GCC	GTC	ACC	GTT	TTG	TAT	GCC	TTC	TGC	ACG	CTC	CTC	1344
Val	Ile	Gly	Val	Ala	Val	Thr	Val	Leu	Tyr	Ala	Phe	Cys	Thr	Leu	Leu	
		435					440					445				
CGC	TGG	AGG	GAC	CCC	GTC	CGT	GGA	CAG	AGC	AGT	GTC	GGC	GTG	GCC	GGA	1392
Arg	Trp	Arg	Asp	Pro	Val	Arg	Gly	Gln	Ser	Ser	Val	Gly	Val	Ala	Gly	
	450					455					460					
GTT	CTG	CTC	ATG	TGC	TTT	AGT	ACC	GCC	GCC	GGA	TTG	GGA	TTG	TCA	GCC	1440
Val	Leu	Leu	Met	Cys	Phe	Ser	Thr	Ala	Ala	Gly	Leu	Gly	Leu	Ser	Ala	
465					470					475					480	
CTG	CTC	GGT	ATC	GTT	TTC	AAT	GCC	GCC	AGC	ACC	CAG	GTG	GTT	CCG	TTT	1488
Leu	Leu	Gly	Ile	Val	Phe	Asn	Ala	Ala	Ser	Thr	Gln	Val	Val	Pro	Phe	
				485					490					495		
TTG	GCC	CTT	GGT	CTG	GGC	GTC	GAT	CAC	ATC	TTC	ATG	CTG	ACC	GCT	GCC	1536
Leu	Ala	Leu	Gly	Leu	Gly	Val	Asp	His	Ile	Phe	Met	Leu	Thr	Ala	Ala	
			500					505						510		

E1
Cont.

TAT	GCG	GAG	AGC	AAT	CGG	CGG	GAG	CAG	ACC	AAG	CTG	ATT	CTC	AAG	AAA	1584
Tyr	Ala	Glu	Ser	Asn	Arg	Arg	Glu	Gln	Thr	Lys	Leu	Ile	Leu	Lys	Lys	
		515					520					525				
GTG	GGA	CCG	AGC	ATC	CTG	TTC	AGT	GCC	TGC	AGC	ACC	GCA	GGA	TCC	TTC	1632
Val	Gly	Pro	Ser	Ile	Leu	Phe	Ser	Ala	Cys	Ser	Thr	Ala	Gly	Ser	Phe	
		530				535					540					
TTT	GCG	GCC	GCC	TTT	ATT	CCG	GTG	CCG	GCT	TTG	AAG	GTA	TTC	TGT	CTG	1680
Phe	Ala	Ala	Ala	Phe	Ile	Pro	Val	Pro	Ala	Leu	Lys	Val	Phe	Cys	Leu	
545					550					555					560	
CAG	GCT	GCC	ATC	GTA	ATG	TGC	TCC	AAT	TTG	GCA	GCG	GCT	CTA	TTG	GTT	1728
Gln	Ala	Ala	Ile	Val	Met	Cys	Ser	Asn	Leu	Ala	Ala	Ala	Leu	Leu	Val	
				565					570					575		
TTT	CCG	GCC	ATG	ATT	TCG	TTG	GAT	CTA	CGG	AGA	CGT	ACC	GCC	GGC	AGG	1776
Phe	Pro	Ala	Met	Ile	Ser	Leu	Asp	Leu	Arg	Arg	Arg	Thr	Ala	Gly	Arg	
			580					585					590			
GCG	GAC	ATC	TTC	TGC	TGC	TGT	TTT	CCG	GTG	TGG	AAG	GAA	CAG	CCG	AAG	1824
Ala	Asp	Ile	Phe	Cys	Cys	Cys	Phe	Pro	Val	Trp	Lys	Glu	Gln	Pro	Lys	
		595					600					605				
GTG	GCA	CCA	CCG	GTG	CTG	CCG	CTG	AAC	AAC	AAC	AAC	GGG	CGC	GGG	GCC	1872
Val	Ala	Pro	Pro	Val	Leu	Pro	Leu	Asn	Asn	Asn	Asn	Gly	Arg	Gly	Ala	
		610				615					620					
CGG	CAT	CCG	AAG	AGC	TGC	AAC	AAC	AAC	AGG	GTG	GCG	CTG	CCC	GCC	CAG	1920
Arg	His	Pro	Lys	Ser	Cys	Asn	Asn	Asn	Arg	Val	Ala	Leu	Pro	Ala	Gln	
625					630					635					640	
AAT	CCT	CTG	CTG	GAA	CAG	AGG	GCA	GAC	ATC	CCT	GGG	AGC	AGT	CAC	TCA	1968
Asn	Pro	Leu	Leu	Glu	Gln	Arg	Ala	Asp	Ile	Pro	Gly	Ser	Ser	His	Ser	
				645					650					655		
CTG	GCG	TCC	TTC	TCT	CTG	GCA	ACA	TTC	GCC	TTT	CAG	CAC	TAC	ACT	CCC	2016
Leu	Ala	Ser	Phe	Ser	Leu	Ala	Thr	Phe	Ala	Phe	Gln	His	Tyr	Thr	Pro	
			660					665					670			
TTC	CTC	ATG	CGC	AGC	TGG	GTG	AAG	TTC	CTG	ACC	GTT	ATG	GGT	TTC	CTG	2064
Phe	Leu	Met	Arg	Ser	Trp	Val	Lys	Phe	Leu	Thr	Val	Met	Gly	Phe	Leu	
		675					680					685				
GCG	GCC	CTC	ATA	TCC	AGC	TTG	TAT	GCC	TCC	ACG	CGC	CTT	CAG	GAT	GGC	2112
Ala	Ala	Leu	Ile	Ser	Ser	Leu	Tyr	Ala	Ser	Thr	Arg	Leu	Gln	Asp	Gly	
		690				695					700					
CTG	GAC	ATT	ATT	GAT	CTG	GTG	CCC	AAG	GAC	AGC	AAC	GAG	CAC	AAG	TTC	2160
Leu	Asp	Ile	Ile	Asp	Leu	Val	Pro	Lys	Asp	Ser	Asn	Glu	His	Lys	Phe	
705					710					715					720	
CTG	GAT	GCT	CAA	ACT	CGG	CTC	TTT	GGC	TTC	TAC	AGC	ATG	TAT	GCG	GTT	2208
Leu	Asp	Ala	Gln	Thr	Arg	Leu	Phe	Gly	Phe	Tyr	Ser	Met	Tyr	Ala	Val	
				725					730					735		

E1
Cont.

ACC	CAG	GGC	AAC	TTT	GAA	TAT	CCC	ACC	CAG	CAG	CAG	TTG	CTC	AGG	GAC	2256
Thr	Gln	Gly	Asn	Phe	Glu	Tyr	Pro	Thr	Gln	Gln	Gln	Leu	Leu	Arg	Asp	
			740					745					750			
TAC	CAT	GAT	TCC	TTT	GTG	CGG	GTG	CCA	CAT	GTG	ATC	AAG	AAT	GAT	AAT	2304
Tyr	His	Asp	Ser	Phe	Val	Arg	Val	Pro	His	Val	Ile	Lys	Asn	Asp	Asn	
		755					760					765				
GGT	GGA	CTG	CCG	GAC	TTC	TGG	CTG	CTG	CTC	TTC	AGC	GAG	TGG	CTG	GGT	2352
Gly	Gly	Leu	Pro	Asp	Phe	Trp	Leu	Leu	Leu	Phe	Ser	Glu	Trp	Leu	Gly	
	770					775					780					
AAT	CTG	CAA	AAG	ATA	TTC	GAC	GAG	GAA	TAC	CGC	GAC	GGA	CGG	CTG	ACC	2400
Asn	Leu	Gln	Lys	Ile	Phe	Asp	Glu	Glu	Tyr	Arg	Asp	Gly	Arg	Leu	Thr	
785					790				795						800	
AAG	GAG	TGC	TGG	TTC	CCA	AAC	GCC	AGC	AGC	GAT	GCC	ATC	CTG	GCC	TAC	2448
Lys	Glu	Cys	Trp	Phe	Pro	Asn	Ala	Ser	Ser	Asp	Ala	Ile	Leu	Ala	Tyr	
				805					810					815		
AAG	CTA	ATC	GTG	CAA	ACC	GGC	CAT	GTG	GAC	AAC	CCC	GTG	GAC	AAG	GAA	2496
Lys	Leu	Ile	Val	Gln	Thr	Gly	His	Val	Asp	Asn	Pro	Val	Asp	Lys	Glu	
			820					825					830			
CTG	GTG	CTC	ACC	AAT	CGC	CTG	GTC	AAC	AGC	GAT	GGC	ATC	ATC	AAC	CAA	2544
Leu	Val	Leu	Thr	Asn	Arg	Leu	Val	Asn	Ser	Asp	Gly	Ile	Ile	Asn	Gln	
		835					840					845				
CGC	GCC	TTC	TAC	AAC	TAT	CTG	TCG	GCA	TGG	GCC	ACC	AAC	GCG	TCT	TCG	2592
Arg	Ala	Phe	Tyr	Asn	Tyr	Leu	Ser	Ala	Trp	Ala	Thr	Asn	Ala	Ser	Ser	
	850					855					860					
CCT	ACG	GAG	CTT	CTC	AGG	GCA	AAT	TGT	ATC	CGG	AAC	CGC	GCC	AAC	GGA	2640
Pro	Thr	Glu	Leu	Leu	Arg	Ala	Asn	Cys	Ile	Arg	Asn	Arg	Ala	Asn	Gly	
865					870					875					880	
GCT	TCT	CAG	GGC	AAA	TTG	TAT	CCG	GAA	CCG	CGC	CAG	TAT	TTT	CAC	CAA	2688
Ala	Ser	Gln	Gly	Lys	Leu	Tyr	Pro	Glu	Pro	Arg	Gln	Tyr	Phe	His	Gln	
				885				890					895			
CCC	AAC	GAG	TAC	GAT	CTT	AAG	ATA	CCC	AAG	AGT	CTG	CCA	TTG	GTC	TAC	2736
Pro	Asn	Glu	Tyr	Asp	Leu	Lys	Ile	Pro	Lys	Ser	Leu	Pro	Leu	Val	Tyr	
			900					905					910			
GCT	CAG	ATG	CCC	TTT	TAC	CTC	CAC	GGA	CTA	ACA	GAT	ACC	TCG	CAG	ATC	2784
Ala	Gln	Met	Pro	Phe	Tyr	Leu	His	Gly	Leu	Thr	Asp	Thr	Ser	Gln	Ile	
		915					920					925				
AAG	ACC	CTG	ATA	GGT	CAT	ATT	CGC	GAC	CTG	AGC	GTC	AAG	TAC	GAG	GGC	2832
Lys	Thr	Leu	Ile	Gly	His	Ile	Arg	Asp	Leu	Ser	Val	Lys	Tyr	Glu	Gly	
	930					935					940					
TTC	GGC	CTG	CCC	AAC	TAT	CCA	TCG	GGC	ATT	CCC	TTC	ATC	TTC	TGG	GAG	2880
Phe	Gly	Leu	Pro	Asn	Tyr	Pro	Ser	Gly	Ile	Pro	Phe	Ile	Phe	Trp	Glu	
945					950					955					960	

E1
Cont.

CAG TAC ATG ACC CTG CGC TCC TCA CTG GCC ATG ATC CTG GCC TGC GTG Gln Tyr Met Thr Leu Arg Ser Ser Leu Ala Met Ile Leu Ala Cys Val 965 970 975	2928
CTA CTC GCC GCC CTG GTG CTG GTC TCC CTG CTC CTG CTC TCC GTT TGG Leu Leu Ala Ala Leu Val Leu Val Ser Leu Leu Leu Leu Ser Val Trp 980 985 990	2976
GCC GCC GTT CTC GTG ATC CTC AGC GTT CTG GCC TCG CTG GCC CAG ATC Ala Ala Val Leu Val Ile Leu Ser Val Leu Ala Ser Leu Ala Gln Ile 995 1000 1005	3024
TTT GGG GCC ATG ACT CTG CTG GGC ATC AAA CTC TCG GCC ATT CCG GCA Phe Gly Ala Met Thr Leu Leu Gly Ile Lys Leu Ser Ala Ile Pro Ala 1010 1015 1020	3072
GTC ATA CTC ATC CTC AGC GTG GGC ATG ATG CTG TGC TTC AAT GTG CTG Val Ile Leu Ile Leu Ser Val Gly Met Met Leu Cys Phe Asn Val Leu 1025 1030 1035 1040	3120
ATA TCA CTG GGC TTC ATG ACA TCC GTT GGC AAC CGA CAG CGC CGC GTC Ile Ser Leu Gly Phe Met Thr Ser Val Gly Asn Arg Gln Arg Arg Val 1045 1050 1055	3168
CAG CTG AGC ATG CAG ATG TCC CTG GGA CCA CTT GTC CAC GGC ATG CTG Gln Leu Ser Met Gln Met Ser Leu Gly Pro Leu Val His Gly Met Leu 1060 1065 1070	3216
ACC TCC GGA GTG GCC GTG TTC ATG CTC TCC ACG TCG CCC TTT GAG TTT Thr Ser Gly Val Ala Val Phe Met Leu Ser Thr Ser Pro Phe Glu Phe 1075 1080 1085	3264
GTG ATC CGG CAC TTC TGC TGG CTT CTG CTG GTG GTC TTA TGC GTT GGC Val Ile Arg His Phe Cys Trp Leu Leu Leu Val Val Leu Cys Val Gly 1090 1095 1100	3312
GCC TGC AAC AGC CTT TTG GTG TTC CCC ATC CTA CTG AGC ATG GTG GGA Ala Cys Asn Ser Leu Leu Val Phe Pro Ile Leu Leu Ser Met Val Gly 1105 1110 1115 1120	3360
CCG GAG GCG GAG CTG GTG CCG CTG GAG CAT CCA GAC CGC ATA TCC ACG Pro Glu Ala Glu Leu Val Pro Leu Glu His Pro Asp Arg Ile Ser Thr 1125 1130 1135	3408
CCC TCT CCG CTG CCC GTG CGC AGC AGC AAG AGA TCG GGC AAA TCC TAT Pro Ser Pro Leu Pro Val Arg Ser Ser Lys Arg Ser Gly Lys Ser Tyr 1140 1145 1150	3456
GTG GTG CAG GGA TCG CGA TCC TCG CGA GGC AGC TGC CAG AAG TCG CAT Val Val Gln Gly Ser Arg Ser Ser Arg Gly Ser Cys Gln Lys Ser His 1155 1160 1165	3504
CAC CAC CAC CAC AAA GAC CTT AAT GAT CCA TCG CTG ACG ACG ATC ACC His His His His Lys Asp Leu Asn Asp Pro Ser Leu Thr Thr Ile Thr 1170 1175 1180	3552

E1
Cont.

190

GAG GAG CCG CAG TCG TGG AAG TCC AGC AAC TCG TCC ATC CAG ATG CCC Glu Glu Pro Gln Ser Trp Lys Ser Ser Asn Ser Ser Ile Gln Met Pro 1185 1190 1195 1200	3600
AAT GAT TGG ACC TAC CAG CCG CGG GAA CAG CGA CCC GCC TCC TAC GCG Asn Asp Trp Thr Tyr Gln Pro Arg Glu Gln Arg Pro Ala Ser Tyr Ala 1205 1210 1215	3648
GCC CCG CCC CCC GCC TAT CAC AAG GCC GCC GCC CAG CAG CAC CAC CAG Ala Pro Pro Pro Ala Tyr His Lys Ala Ala Ala Gln Gln His His Gln 1220 1225 1230	3696
CAT CAG GGC CCG CCC ACA ACG CCC CCG CCG CCC TTC CCG ACG GCC TAT His Gln Gly Pro Pro Thr Thr Pro Pro Pro Pro Phe Pro Thr Ala Tyr 1235 1240 1245	3744
CCG CCG GAG CTG CAG AGC ATC GTG GTG CAG CCG GAG GTG ACG GTG GAG Pro Pro Glu Leu Gln Ser Ile Val Val Gln Pro Glu Val Thr Val Glu 1250 1255 1260	3792
ACG ACG CAC TCG GAC AGC AAC ACC ACC AAG GTG ACG GCC ACG GCC AAC Thr Thr His Ser Asp Ser Asn Thr Thr Lys Val Thr Ala Thr Ala Asn 1265 1270 1275 1280	3840
ATC AAG GTG GAG CTG GCC ATG CCC GGC AGG GCG GTG CGC AGC TAT AAC Ile Lys Val Glu Leu Ala Met Pro Gly Arg Ala Val Arg Ser Tyr Asn 1285 1290 1295	3888
TTT ACG AGT TAG Phe Thr Ser	3900

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ACCGAGGGCT GGGACGAAGA TGGC

24

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

*E1
Cont.*

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met Gly Ser Ser His His His His His His Leu Val Pro Arg Gly Ser
1 5 10 15

E!
Cont.

His Met

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1299 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Asp Arg Asp Ser Leu Pro Arg Val Pro Asp Thr His Gly Asp Val
 1 5 10 15
 Val Asp Glu Lys Leu Phe Ser Asp Leu Tyr Ile Arg Thr Ser Trp Val
 20 25 30
 Asp Ala Gln Val Ala Leu Asp Gln Ile Asp Lys Gly Lys Ala Arg Gly
 35 40 45
 Ser Arg Thr Ala Ile Tyr Leu Arg Ser Val Phe Gln Ser His Leu Glu
 50 55 60
 Thr Leu Gly Ser Ser Val Gln Lys His Ala Gly Lys Val Leu Phe Val
 65 70 75 80
 Ala Ile Leu Val Leu Ser Thr Phe Cys Val Gly Leu Lys Ser Ala Gln
 85 90 95
 Ile His Ser Lys Val His Gln Leu Trp Ile Gln Glu Gly Gly Gly Leu
 100 105 110
 Glu Ala Glu Leu Ala Tyr Thr Gln Lys Thr Ile Gly Glu Asp Glu Ser
 115 120 125
 Ala Thr His Gln Leu Leu Ile Gln Thr Thr His Asp Pro Asn Ala Ser
 130 135 140
 Val Leu His Pro Gln Ala Leu Leu Ala His Leu Glu Val Leu Val Lys
 145 150 155 160
 Ala Thr Ala Val Lys Val His Leu Tyr Asp Thr Glu Trp Gly Leu Arg
 165 170 175
 Asp Met Cys Asn Met Pro Ser Thr Pro Ser Phe Glu Gly Ile Tyr Tyr
 180 185 190
 Ile Glu Gln Ile Leu Arg His Leu Ile Pro Cys Ser Ile Ile Thr Pro
 195 200 205
 Leu Asp Cys Phe Trp Glu Gly Ser Gln Leu Leu Gly Pro Glu Ser Ala
 210 215 220

E1
Cont

Val Val Ile Pro Gly Leu Asn Gln Arg Leu Leu Trp Thr Thr Leu Asn
 225 230 235 240
 Pro Ala Ser Val Met Gln Tyr Met Lys Gln Lys Met Ser Glu Glu Lys
 245 250 255
 Ile Ser Phe Asp Phe Glu Thr Val Glu Gln Tyr Met Lys Arg Ala Ala
 260 265 270
 Ile Ala Ser Gly Tyr Met Glu Lys Pro Cys Leu Asn Pro Leu Asn Pro
 275 280 285
 Asn Cys Pro Asp Thr Ala Pro Asn Lys Asn Ser Thr Gln Pro Pro Asp
 290 295 300
 Val Gly Ala Ile Leu Ser Gly Gly Cys Tyr Gly Tyr Ala Ala Lys His
 305 310 315 320
 Met His Trp Pro Glu Glu Leu Ile Val Gly Gly Ala Lys Arg Asn Arg
 325 330 335
 Ser Gly His Leu Arg Lys Ala Gln Ala Leu Gln Ser Val Val Gln Leu
 340 345 350
 Met Thr Glu Lys Glu Met Tyr Asp Gln Trp Gln Asp Asn Tyr Lys Val
 355 360 365
 His His Leu Gly Trp Thr Gln Glu Lys Ala Ala Glu Val Leu Asn Ala
 370 375 380
 Trp Gln Arg Asn Phe Ser Arg Glu Val Glu Gln Leu Leu Arg Lys Gln
 385 390 395 400
 Ser Arg Ile Ala Thr Asn Tyr Asp Ile Tyr Val Phe Ser Ser Ala Ala
 405 410 415
 Leu Asp Asp Ile Leu Ala Lys Phe Ser His Pro Ser Ala Leu Ser Ile
 420 425 430
 Val Ile Gly Val Ala Val Thr Val Leu Tyr Ala Phe Cys Thr Leu Leu
 435 440 445
 Arg Trp Arg Asp Pro Val Arg Gly Gln Ser Ser Val Gly Val Ala Gly
 450 455 460
 Val Leu Leu Met Cys Phe Ser Thr Ala Ala Gly Leu Gly Leu Ser Ala
 465 470 475 480
 Leu Leu Gly Ile Val Phe Asn Ala Ala Ser Thr Gln Val Val Pro Phe
 485 490 495
 Leu Ala Leu Gly Leu Gly Val Asp His Ile Phe Met Leu Thr Ala Ala
 500 505 510
 Tyr Ala Glu Ser Asn Arg Arg Glu Gln Thr Lys Leu Ile Leu Lys Lys
 515 520 525

E!
 Cont

Val Gly Pro Ser Ile Leu Phe Ser Ala Cys Ser Thr Ala Gly Ser Phe
 530 535 540
 Phe Ala Ala Ala Phe Ile Pro Val Pro Ala Leu Lys Val Phe Cys Leu
 545 550 555 560
 Gln Ala Ala Ile Val Met Cys Ser Asn Leu Ala Ala Ala Leu Leu Val
 565 570 575
 Phe Pro Ala Met Ile Ser Leu Asp Leu Arg Arg Arg Thr Ala Gly Arg
 580 585 590
 Ala Asp Ile Phe Cys Cys Cys Phe Pro Val Trp Lys Glu Gln Pro Lys
 595 600 605
 Val Ala Pro Pro Val Leu Pro Leu Asn Asn Asn Asn Gly Arg Gly Ala
 610 615 620
 Arg His Pro Lys Ser Cys Asn Asn Asn Arg Val Ala Leu Pro Ala Gln
 625 630 635 640
 Asn Pro Leu Leu Glu Gln Arg Ala Asp Ile Pro Gly Ser Ser His Ser
 645 650 655
 Leu Ala Ser Phe Ser Leu Ala Thr Phe Ala Phe Gln His Tyr Thr Pro
 660 665 670
 Phe Leu Met Arg Ser Trp Val Lys Phe Leu Thr Val Met Gly Phe Leu
 675 680 685
 Ala Ala Leu Ile Ser Ser Leu Tyr Ala Ser Thr Arg Leu Gln Asp Gly
 690 695 700
 Leu Asp Ile Ile Asp Leu Val Pro Lys Asp Ser Asn Glu His Lys Phe
 705 710 715 720
 Leu Asp Ala Gln Thr Arg Leu Phe Gly Phe Tyr Ser Met Tyr Ala Val
 725 730 735
 Thr Gln Gly Asn Phe Glu Tyr Pro Thr Gln Gln Gln Leu Leu Arg Asp
 740 745 750
 Tyr His Asp Ser Phe Val Arg Val Pro His Val Ile Lys Asn Asp Asn
 755 760 765
 Gly Gly Leu Pro Asp Phe Trp Leu Leu Leu Phe Ser Glu Trp Leu Gly
 770 775 780
 Asn Leu Gln Lys Ile Phe Asp Glu Glu Tyr Arg Asp Gly Arg Leu Thr
 785 790 795 800
 Lys Glu Cys Trp Phe Pro Asn Ala Ser Ser Asp Ala Ile Leu Ala Tyr
 805 810 815
 Lys Leu Ile Val Gln Thr Gly His Val Asp Asn Pro Val Asp Lys Glu
 820 825 830

E1
 Cont.

Leu Val Leu Thr Asn Arg Leu Val Asn Ser Asp Gly Ile Ile Asn Gln
 835 840 845
 Arg Ala Phe Tyr Asn Tyr Leu Ser Ala Trp Ala Thr Asn Ala Ser Ser
 850 855 860
 Pro Thr Glu Leu Leu Arg Ala Asn Cys Ile Arg Asn Arg Ala Asn Gly
 865 870 875 880
 Ala Ser Gln Gly Lys Leu Tyr Pro Glu Pro Arg Gln Tyr Phe His Gln
 885 890 895
 Pro Asn Glu Tyr Asp Leu Lys Ile Pro Lys Ser Leu Pro Leu Val Tyr
 900 905 910
 Ala Gln Met Pro Phe Tyr Leu His Gly Leu Thr Asp Thr Ser Gln Ile
 915 920 925
 Lys Thr Leu Ile Gly His Ile Arg Asp Leu Ser Val Lys Tyr Glu Gly
 930 935 940
 Phe Gly Leu Pro Asn Tyr Pro Ser Gly Ile Pro Phe Ile Phe Trp Glu
 945 950 955 960
 Gln Tyr Met Thr Leu Arg Ser Ser Leu Ala Met Ile Leu Ala Cys Val
 965 970 975
 Leu Leu Ala Ala Leu Val Leu Val Ser Leu Leu Leu Leu Ser Val Trp
 980 985 990
 Ala Ala Val Leu Val Ile Leu Ser Val Leu Ala Ser Leu Ala Gln Ile
 995 1000 1005
 Phe Gly Ala Met Thr Leu Leu Gly Ile Lys Leu Ser Ala Ile Pro Ala
 1010 1015 1020
 Val Ile Leu Ile Leu Ser Val Gly Met Met Leu Cys Phe Asn Val Leu
 1025 1030 1035 1040
 Ile Ser Leu Gly Phe Met Thr Ser Val Gly Asn Arg Gln Arg Arg Val
 1045 1050 1055
 Gln Leu Ser Met Gln Met Ser Leu Gly Pro Leu Val His Gly Met Leu
 1060 1065 1070
 Thr Ser Gly Val Ala Val Phe Met Leu Ser Thr Ser Pro Phe Glu Phe
 1075 1080 1085
 Val Ile Arg His Phe Cys Trp Leu Leu Leu Val Val Leu Cys Val Gly
 1090 1095 1100
 Ala Cys Asn Ser Leu Leu Val Phe Pro Ile Leu Leu Ser Met Val Gly
 1105 1110 1115 1120
 Pro Glu Ala Glu Leu Val Pro Leu Glu His Pro Asp Arg Ile Ser Thr
 1125 1130 1135

E1
 Cont.

Pro Ser Pro Leu Pro Val Arg Ser Ser Lys Arg Ser Gly Lys Ser Tyr
 1140 1145 1150
 Val Val Gln Gly Ser Arg Ser Ser Arg Gly Ser Cys Gln Lys Ser His
 1155 1160 1165
 His His His His Lys Asp Leu Asn Asp Pro Ser Leu Thr Thr Ile Thr
 1170 1175 1180
 Glu Glu Pro Gln Ser Trp Lys Ser Ser Asn Ser Ser Ile Gln Met Pro
 1185 1190 1195 1200
 Asn Asp Trp Thr Tyr Gln Pro Arg Glu Gln Arg Pro Ala Ser Tyr Ala
 1205 1210 1215
 Ala Pro Pro Pro Ala Tyr His Lys Ala Ala Ala Gln Gln His His Gln
 1220 1225 1230
 His Gln Gly Pro Pro Thr Thr Pro Pro Pro Pro Phe Pro Thr Ala Tyr
 1235 1240 1245
 Pro Pro Glu Leu Gln Ser Ile Val Val Gln Pro Glu Val Thr Val Glu
 1250 1255 1260
 Thr Thr His Ser Asp Ser Asn Thr Thr Lys Val Thr Ala Thr Ala Asn
 1265 1270 1275 1280
 Ile Lys Val Glu Leu Ala Met Pro Gly Arg Ala Val Arg Ser Tyr Asn
 1285 1290 1295
 Phe Thr Ser

El
 Cont